National Defence / Defense nationale

UNCLASSIFIED

DRES

SUFFIELD REPORT

NO. 549

(1)

A FORMULATION OF A STOCHASTIC SAMPLING ERROR MODEL AND A SIGNAL DETECTION ALGORITHM FOR THE AERODYNAMIC PARTICLE SIZE ANALYSER

by

E. Yee

PCN 351 SP

June 1991

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD, RALSTON, ALBERTA

Canada

UNCLASSIFIED

DEFENCE RESEARCH ESTABLISHMENT SUFFIELD

RALSTON ALBERTA

SUFFIELD REPORT NO. 549

A FORMULATION OF A STOCHASTIC SAMPLING ERROR MODEL AND A SIGNAL DETECTION ALGORITHM FOR THE AERODYNAMIC PARTICLE SIZE ANALYZER

by

E. Yee

PCN No. 351SP

UNCLASSIFIED

ABSTRACT

The determination of the distribution of airborne toxic particles as a function of the aerodynamic diameter provides important information as well as criteria for the definition of hazard as applied to levels of airborne contamination. This is because the aerodynamic particle size distribution embodies the information related to particle density, diameter, shape factor and slip correction that is critical for the characterization of particle motion in settling and impaction and it is these motions that are responsible for particle deposition in the respiratory tract and particle collection in aerosol sampling devices. For a given definition of hazard based on some parameter related to the aerodynamic size distribution, this paper develops a statistical sampling error model for the parameter that is based on the Poisson process. Given that an appropriate sampling program has been designed for the measurement of the size distribution-related parameter with the aerodynamic particle size analyzer, this paper proceeds to the derivation of an optimum detection algorithm for the detection of a signal aerosol sequence in a set of J aerosol samples with a common background. The detection algorithm is based on the generalized likelihood ratio test in which the received count associated with the aerosol sample is modeled as a Poisson distributed random variable. Performance analyses of the resulting algorithm, based on the probability of detection (P_D) versus signal-to-noise ratio for several given fixed false alarm probabilities (P_{FA}), are presented.

RÉSUMÉ

La détermination de la distribution des particules en suspension dans l'air en fonction de leur diamètre aérodynamique fournit d'importantes informations ainsi que des critères qui permettent de définir les risques correspondant aux degrés de contamination du milieu atmosphérique. En effet, la distribution de la taille aérodynamique renferme des informations sur la masse volumique, le diamètre, le facteur de forme et la correction de glissement des particules, qui sont essentielles pour caractériser le mouvement des particules au cours de leur dépôt et de leur impact; ce mouvement est également responsable du dépôt des particules dans les voies respiratoires et de leur piégeage dans les dispositifs de prélèvement d'aérosols. Pour une définition donnée du risque déterminé à partir d'un paramètre quelconque lié à la distribution de la taille aérodynamique, on élabore dans cette communication un modèle statistique basé sur la loi de Poisson, permettant de déterminer l'erreur d'échantillonnage pour le paramètre. À l'aide d'un programme de prélèvement approprié conçu pour mesurer, avec l'analyseur de tailles aérodynamiques, le paramètre lié à distribution de la taille, on déduit, dans cette communication, l'algorithme optimal permettant de détecter une séquence d'aérosols dans un ensemble de J échantillons d'aérosols qui présentent les mêmes caractéristiques de base. L'algorithme de détection est basé sur un test généralisé du rapport des vraisemblances, dans lequel la valeur reçue associée à l'échantillon d'aérosol est modélisée comme une variable aléatoire qui obéit à la loi de Poisson. On présente les résultats des analyses de performance de l'algorithme ainsi obtenu, basés sur la probabilité de détection (P_D) en fonction du rapport signal/bruit pour plusieurs valeurs données de probabilités de fausse alarme (P_{FA}).

TABLE OF CONTENTS

I. INTRODUCTION

The aerodynamic particle sizer (APS), Model 3300 (TSI Incorporated) can provide a real-time measurement of the particle size distribution according to the aerodynamic diameter [1]. This sampling instrument determines the aerodynamic diameter of the individual aerosol particles in the sample by measuring the transit time of the particles between two spots generated by a laser velocimeter that employs a 2 mW polarized He-Ne laser as the light source.

The APS consists of two basic modules: (1) the sensor module and (2) the signal processing and microcomputer module. The main parts of the sensor module are the accelerating nozzle and the two spot laser velocimeter. This module brings the aerosol sample into an outer accelerating orifice and focusses the individual particles through a dual beam laser formed by splitting a focussed laser beam, on the basis of polarization, using a calcite plate. The beams are then focussed using a cylindrical lens to produce two flat beams of rectangular cross-section just downstream of the nozzle orifice. As the aerosol particle passes through these two beams, it triggers a pair of electrical pulses whose temporal separation is accurately measured using a high speed digital clock. The conversion of this information into an aerodynamic particle size is directed by the signal pro essing module. A multichannel accumulator (MCA) is used to record the transit times of all aerosol particles and, at the end of the prescribed sampling period, a microcomputer reads each channel of the MCA, translates the channel numbers to aerodynamic particle sizes, and displays the information as a discrete spectrum or histogram consisting of 48 size intervals (i.e., bins) spanning the aerosol diameter range from 0.5–15 μm.

Currently, there is active interest in the development of signal processing algorithms for the detection of weak aerosol target signals in a heavy clutter aerosol background using an APS as the sensor detection system. A theoretical investigation of the performance characteristics of the APS, when utilized in the detection mode, is undertaken in this paper. Towards this objective, a stochastic sampling error model and a signal detection algorithm based on the application of the classical generalized maximum likelihood ratio test, are developed for the APS. This paper is organized as follows. In Section II, a statistical sampling error model for the APS, based on the Poisson process, is developed and analyzed. A signal detection algorithm which applies the generalized maximum likelihood ratio test to the Poisson process is developed in Section III. In order to analyze the performance of the detection algorithm, probabilities of false alarm and detection as a function of

the aerosol target-to-clutter signal-to-noise ratio (SNR) are also presented in Section III. Section IV provides an application of the statistical sampling error model and the signal detection algorithm to two explicit numerical examples involving the detection of a transient aerosol target signal with an unknown arrival time in a background clutter aerosol environment. Finally, conclusions are listed in Section V.

II. A STATISTICAL SAMPLING ERROR MODEL FOR THE APS

The measurement of the aerodynamic particle size distribution function (PSDF) or, for that matter, any functional of the PSDF is necessarily subject to an intrinsic variability that is due both to statistical sampling errors and to natural physical variations which are not readily separable from the statistical errors. For a single sampling instrument such as the APS, the statistical sampling errors can be reduced by the design of an appropriate sampling method. This section is concerned with the estimation of the sampling errors associated with the operation of the APS with the objective of specifying the required sample size and observation times that must be adopted in order to bound such errors.

The sampling error model for the APS is based on the Poisson model which provides the information about the statistical properties of the random point process associated with the fluctuations of any size distribution-related variable as a consequence of the stochastic fluctuations of aerosol particle numbers in a given size range in the sampling volume. It is asserted that the particle number fluctuations in a given aerodynamic diameter range can be adequately modeled by a Poisson random variable [2] with the probability $k = 0, 1, 2, \ldots$ particles occurring given by

$$P_k = \frac{\lambda^k}{k!} \exp(-\lambda), \qquad \lambda \geq 0,$$

where λ is the Poisson parameter.

With the assumption that the Poisson distribution properly describes the random fluctuations of aerosol particle numbers of a given aerodynamic diameter range in the sampling volume V_s, the variability of any size distribution-related variable can be derived as follows. For any given function $f(D)$ of the aerodynamic diameter D, any particular measurement of the observed value of the function due to aerosol particles in the size interval D to $D + dD$ in the sampled volume V_s, is given by $dF = f(D)n(D)\,V_s\,dD$, where $n(D)$ is a particular realization of the aerodynamic particle size distribution. Consequently, it can be concluded that the ensemble mean of the stochastic functional F, due to the contribution of all aerosol particles within the aerodynamic diameter range

$[D_1, D_2]$ in the sampling volume V_s, is given by

$$\langle F \rangle = \int_{D_1}^{D_2} f(D)\langle n(D)\rangle\, V_s\, dD = \int_{D_1}^{D_2} f(D)N(D)\, V_s\, dD, \tag{1}$$

where the ensemble average of any quantity is indicated by angled brackets, $\langle\ \rangle$. It is worthwhile pointing out that the ensemble mean number of aerosol particles with diameters D to $D + dD$ in volume V_s is given by $N(D)\, V_s dD \equiv \langle n(D)\rangle\, V_s\, dD$. In other words, the aerosol particles sampled in the volume V_s are Poisson distributed about the ensemble mean particle size distribution, $N(D)$. Similarly, the variance of F can be written as

$$\mathrm{Var}\big(F\big) \equiv \sigma_F^2 = \int_{D_1}^{D_2} f^2(D)\mathrm{Var}\big(n(D)\big)\, V_s\, dD.$$

Now, recall that the mean and variance of a Poisson distribution are equal. Since it has been assumed that $n(D)$ at any fixed value of D is a Poisson random variable, it follows that $\mathrm{Var}\big(n(D)\big) \equiv \sigma^2_{n(D)} = \langle n(D)\rangle = N(D)$ so

$$\sigma_F^2 = \int_{D_1}^{D_2} f^2(D)N(D)\, V_s\, dD. \tag{2}$$

In light of Eqs. (1) and (2), the relative fluctuation intensity of F defined as $I_F \equiv \sigma_F/\langle F\rangle$, can be expressed as follows:

$$I_F \equiv \frac{\sigma_F}{\langle F\rangle} = \frac{\left(\int_{D_1}^{D_2} f^2(D)N(D)\, V_s\, dD\right)^{1/2}}{\int_{D_1}^{D_2} f(D)N(D)\, V_s\, dD}. \tag{3}$$

Since several particulate properties are proportional to the diameter of the particles raised to an appropriate power, it is useful to consider functions $f(D)$ with the form $f(D) = c_n D^n$, where c_n is a constant shape factor whose precise form is determined only by the specific value of n. Examples of such functions include the following: $f(D) = 1$ (i.e., $n = 0$ and $c_0 = 1$) for which the associated property is the expected value (ensemble average) of the total number of aerosol particles within the size interval $[D_1, D_2]$ in the sampled volume; $f(D) = \pi D^2$ (i.e., $n = 2$ and $c_2 = \pi$) for which the associated property is the equivalent surface area of the aerosol particles; and, $f(D) = \pi\rho D^3/6$ (i.e., $n = 3$ and $c_3 = \pi\rho/6$ where ρ is the particle density) for which the associated property is the mass of the aerosol particles. Insertion of this specialized form for $f(D)$ into Eq. (3) yields

$$I_{F_n} \equiv \frac{\sigma_{F_n}}{\langle F_n\rangle} = \frac{\left(\int_{D_1}^{D_2} D^{2n}N(D)\, V_s\, dD\right)^{1/2}}{\int_{D_1}^{D_2} D^n N(D)\, V_s\, dD}. \tag{4}$$

Since the mth moment of the particle diameter about the origin $D = 0$ over the size interval $[D_1, D_2]$ is defined as

$$\langle D^m \rangle \equiv \frac{\int_{D_1}^{D_2} D^m N(D)\, dD}{\int_{D_1}^{D_2} N(D)\, dD},$$

it follows easily that the relative fluctuation intensity (cf. Eq. (4)) can be written in the following form:

$$I_{F_n} \equiv \frac{\sigma_{F_n}}{\langle F_n \rangle} = \langle N \rangle^{-1/2} \frac{\langle D^{2n} \rangle^{1/2}}{\langle D^n \rangle}, \tag{5}$$

where $\langle N \rangle$, the expected total number of particles in the size diameter range $[D_1, D_2]$ in the sampled volume V_s, is given by

$$\langle N \rangle = \int_{D_1}^{D_2} N(D)\, V_s\, dD.$$

It is informative to write Eq. (5) as

$$I_{F_n} \equiv \frac{\sigma_{F_n}}{\langle F_n \rangle} = N_e^{-1/2}, \tag{6a}$$

where N_e is an effective aerosol particle number which is defined by

$$N_e = \kappa \langle N \rangle \tag{6b}$$

and

$$\kappa = \frac{\langle D^n \rangle^2}{\langle D^{2n} \rangle}. \tag{6c}$$

In light of this development, it is seen that the relative fluctuation intensity of any particle size distribution-related variable F_n is proportional to the inverse square root of the effective total number of aerosol particles, N_e. The effective total number of particles, N_e, is directly proportional to the expected value of the total number of aerosol particles, $\langle N \rangle$, with the constant of proportionality, κ, involving the moments of the particle size distribution as per Eq. (6c).

In order to provide some concrete results concerning the statistical sampling error model derived above, it is useful at this point to choose an explicit form for the aerodynamic particle size distribution function, $N(D)$. For this purpose, it is assumed that the aerosol particle size distribution can be adequately approximated by a three parameter gamma size distribution represented by

$$N(D) = \frac{N^*}{\Gamma(\alpha)\beta^\alpha} \exp(-D/\beta) D^{\alpha-1},$$

where $\alpha > 0$, $\beta > 0$, $D \geq 0$ and

$$\Gamma(\alpha) \equiv \int_0^\infty t^{\alpha-1} \exp(t)\, dt$$

is the gamma function. Here, α is the shape parameter, β the scale parameter, and N^* the particle number density. It is noted that the gamma distribution provides an attractive compromise between the widely used log-normal and exponential distributions that has been traditionally assumed for the aerosol size distribution. Indeed, measurements of Berger, Melice and Demuth [3] to test the applicability of the log-normal and gamma distribution on an extensive data set for suspended particulates in Gent, Belgium have demonstrated conclusively that the gamma distribution was superior to the log-normal distribution for modeling the suspended particulate levels. For the gamma size distribution, $N(D)$, the relative fluctuation intensity I_{F_n} expressed in Eq. (4) reduces to the following form:

$$I_{F_n} \equiv \frac{\sigma_{F_n}}{\langle F_n \rangle} = \frac{1}{\sqrt{N^* V_s}} \frac{\left[P(2n+\alpha, D_2/\beta) - P(2n+\alpha, D_1/\beta)\right]^{1/2}}{P(n+\alpha, D_2/\beta) - P(n+\alpha, D_1/\beta)}, \tag{7}$$

where $P(\alpha, x)$ is the incomplete gamma function defined as

$$P(\alpha, x) \equiv \frac{\gamma(\alpha, x)}{\Gamma(\alpha)} = \frac{1}{\Gamma(\alpha)} \int_0^x t^{\alpha-1} \exp(-t)\, dt.$$

Similarly, it can be readily shown that the parameter κ for the gamma size distribution can be represented by

$$\kappa = \frac{\left[P(n+\alpha, D_2/\beta) - P(n+\alpha, D_1/\beta)\right]^2}{\left[P(\alpha, D_2/\beta) - P(\alpha, D_1/\beta)\right]\left[P(2n+\alpha, D_2/\beta) - P(2n+\alpha, D_1/\beta)\right]}. \tag{8}$$

For a fixed size interval $[D_1, D_2]$, the parameter κ is a function only of n, the exponent related to the size distribution-related variable F_n and of α and β, the gamma shape and scale parameters, respectively. A plot of $k \equiv 1/\kappa = \langle N \rangle / N_c$ as a function of n for $\alpha = 1, 2$ and $\beta = 1.0$ μm, $D_1 = 0.5$ μm and $D_2 = 15.0$ μm is shown in Figure 1. It is evident from this plot that as the exponent n increases, thereby providing increasing weight to the aerosol particles with the larger diameters, κ (and, therefore the effective particle number, N_c, as per Eq. (6b)) decreases very rapidly, thus increasing the relative fluctuation intensity of F_n. Intuitively, this is due to the smaller number of aerosol particles which contribute to the major fraction of F_n as n increases. Furthermore, note that $\kappa = 1$ when $n = 0$, with the result that in this case, Eq. (6a) reduces to $\sigma_N / \langle N \rangle = \langle N \rangle^{-1/2}$. The latter relationship is a well-known result for the relative fluctuation intensity of the total number of aerosol particles, N, in the sampled volume V_s.

By application of Eq. (7), it is possible to construct curves for the normalized relative fluctuation intensity for various size distribution-related variables. Towards this end, we consider the size interval $[D_1, D_2] = [0.5, 15.0]$ μm which is the operating particle size range for the APS.

With this operating constraint, curves for the normalized relative fluctuation intensity, defined as $(N^* V_s)^{1/2} I_{F_n}$, have been calculated as a function of the gamma size distribution shape parameter α and are displayed in Figures 2, 3, and 4 for $\beta = 1.0$ μm and $n = 0, 2, 3$, respectively. These curves distinctly show the variation of the normalized relative fluctuation intensity of F_n with the shape of the underlying number size distribution function as embodied in the gamma shape parameter α. To provide further illustration of the effects of the number size distribution shape on the relative fluctuation intensity, we consider the ratio, R, defined as the ratio of I_{F_n} using the gamma size distribution to that using the exponential size distribution (to which the gamma distribution converges as $\alpha \rightarrow 1$). In light of Eq. (7), the ratio R can be expressed as

$$R \equiv \frac{I_{F_n}(\text{gamma})}{I_{F_n}(\text{exponential})} = \frac{\left[P(2n+\alpha, D_2/\beta) - P(2n+\alpha, D_1/\beta)\right]^{1/2}\left[P(n+1, D_2/\beta) - P(n+1, D_1/\beta)\right]}{\left[P(2n+1, D_2/\beta) - P(2n+1, D_1/\beta)\right]^{1/2}\left[P(n+\alpha, D_2/\beta) - P(n+\alpha, D_1/\beta)\right]}.$$

Plots of the ratio R with changing α for the cases of n equal to 0, 2 and 3 are displayed in Figure 5 for $\beta = 1.0$ μm and in Figure 6 for $\beta = 3.0$ μm. These plots provide a visual indication of the distribution shape effects on the relative fluctuation intensity of the sampled variable F_n, based on differences between the two sampling error models. The most important point to note in these plots is that there can be a change in the relative fluctuation intensity of the variable F_n that arises simply because of a variation in the size distribution shape. For example, when a gamma size distribution having an actual shape parameter $\alpha = 0.1$ ($\beta = 1.0$ μm) is sampled with an APS and the relative fluctuation intensity of the number concentration (i.e., $n = 0$) in the diameter range $[D_1, D_2] = [0.5, 15.0]$ μm is subsequently computed using an exponential model (i.e., $\alpha \rightarrow 1$), then the actual relative fluctuation intensity of the number concentration in this case is about 5 times the one computed. This implies that there is an underestimation by about 5 times the actual relative fluctuation intensity value, which can arise simply by neglecting the effects of the distribution shape.

As a final point, it should be noted that the statistical sampling error model developed in this section can be used to compute the sampling errors for a large number of variables related to the aerodynamic particle size distribution. In particular, the model can be applied to determine the minimum sampling duration, t_s, that would be required for operation of the APS at a given aerosol sampling flowrate, Q_s, in order that the sampling error of a given variable of interest (e.g., F_n) be smaller than or equal to some prescribed I_{F_n}. As a simple example, to achieve a sampling error of one percent for the total aerosol particle number using the APS with $D_1 = 0.5$ μm and $D_2 = 15.0$ μm for an underlying size distribution function with $\alpha = 1.5$, $\beta = 2.5$ μm, and $N^* = 6.0$ cm^{-3} will

require a minimum sampling volume of $V_s = 1800$ cm^3 (cf. Eq. (7)). For the APS operating at an aerosol flowrate of $Q_s = 83.3$ cm^3 s^{-1}, this requires a sampling duration of $t_s = V_s/Q_s = 22$ s.

III. DETECTION ALGORITHM AND PERFORMANCE ANALYSIS

In this section, a detector based on the parameter F_n, whose statistical sampling errors were analyzed in the previous section, is proposed. Towards this objective, consider the APS sampling for duration t_s at an aerosol flowrate Q_s to produce count data as the set of numbers $\{F_n(j), j = 1, 2, \ldots, J\}$, viz., these numbers are the observed values of the parameter F_n due to aerosol particles in the size interval $[D_1, D_2]$ in the sampled volume $V_s = Q_s t_s$. These counts are random variables and, as in the previous section, it is assumed that the count $F_n(j)$ during the jth sampling period (reception or measurement) is a Poisson random variable that assumes the values $k = 0, 1, 2, \ldots$ with probability $P_k = \left(\lambda^k \exp(-\lambda)\right)/k!$ $(\lambda > 0)$. The counts obtained during disjoint sampling periods (i.e., exposures) are assumed to be independent random variables.

For the purposes of detection, it is assumed that the aerosol sample is the sum of a target "signal" aerosol whose presence is to be detected and a background "clutter" aerosol whose presence is unavoidably received from the natural aerosol component present in the atmosphere. In this case, each count, $F_n(j)$, will be a Poisson random variable with parameter $\lambda^{(1)}(j) = S(j) + \lambda^{(0)}(j)$, where $\lambda^{(1)}(j)$ refers to the situation when the signal aerosol to be detected is actually present in the sample, a situation that is referred to as hypothesis H_1. When the signal aerosol is absent in the sample—a situation referred to as hypothesis H_0—the count data $F_n(j)$ will still be a Poisson random variable, but with parameter $\lambda^{(0)}(j)$, where $\lambda^{(0)}(j)$ refers only to the background aerosol component. In the ensuing analysis, it is assumed that $\{\lambda^{(0)}(j)\}$ do not depend on j, viz., the background Poisson parameter is the same for all J aerosol samples obtained. Furthermore, it is assumed that $\{S(j)\}$ is an *a priori* known, non-negative sequence due to the presence of the target aerosol signal.

The observed count data $\{F_n(j), j = 1, 2, \ldots, J\}$ is a set of mutually independent Poisson random variables with parameters $\lambda^{(1)}(j) = S(j) + \lambda^{(0)}$ if the signal aerosol is present (hypothesis H_1) or, with parameters $\lambda^{(0)}(j) = \lambda^{(0)}$ if the signal aerosol is absent (hypothesis H_0). This forms the basis for the conventional binary hypothesis testing paradigm in which the system receives a sequence of aerosol sample count data $F_n(j)$ from J independent sampling periods that can correspond to a Poisson process with parameter $\lambda(j)$ such that

$$H_0 : \lambda(j) = \lambda^{(0)}, \qquad j = 1, 2, \ldots, J;$$

$$H_1 : \lambda(j) = S(j) + \lambda^{(0)}, \qquad j = 1, 2, \ldots, J.$$

Detection of the signal aerosol can then be formulated as testing the null hypothesis H_0 versus the alternative hypothesis H_1. This objective can be achieved with the likelihood ratio test whose purpose is to test hypothesis H_1 against H_0, holding the probability that H_1 is accepted when H_0 is actually true, i.e., the probability of false alarm, below some fixed error level.

The detector is based on a generalized likelihood ratio test and is localized, viz., it examines a set of count data spanning a relatively short sampling duration. While not structurally necessary, localized detection is better suited to the nature of the signals that are assumed to be present (time-varying and transient) than a global detection. The likelihood ratio Λ is the ratio of the probability ("likelihood") $p(\{F_n(j)\}|H_1)$ of observing the count data $\{F_n(j), j = 1, 2, \ldots, J\}$ under H_1 to the probability of observing the data under H_0, $p(\{F_n(j)\}|H_0)$ so

$$\Lambda = \frac{p(\{F_n(j)\}|H_1)}{p(\{F_n(j)\}|H_0)}.$$

If $\lambda^{(0)}$ is known, then as already noted, $\{F_n(j)\}$ are independent, Poisson random variables, so the joint probability $p(\{F_n(j)\}|H_i)$ $(i = 0, 1)$ is the J-fold product of the individual probability $p(F_n(j)|H_i)$ of observing each $F_n(j)$. Consequently,

$$p(\{F_n(j)\}|H_i) = \prod_{j=1}^{J} \frac{[\lambda^{(i)}(j)]^{F_n(j)}}{F_n(j)!} \exp(-\lambda^{(i)}(j))$$

and so it follows

$$\Lambda = \frac{\prod_{j=1}^{J} [S(j) + \lambda^{(0)}]^{F_n(j)} \exp(-(S(j) + \lambda^{(0)})) / F_n(j)!}{\prod_{j=1}^{J} [\lambda^{(0)}]^{F_n(j)} \exp(-\lambda^{(0)}) / F_n(j)!}.$$

Some straightforward algebraic manipulations can be performed to simplify Λ to the following form:

$$\Lambda = \exp\left(-\sum_{j=1}^{J} S(j)\right) \prod_{j=1}^{J} \left(1 + \frac{S(j)}{\lambda^{(0)}}\right)^{F_n(j)}.$$

The likelihood ratio must now be compared against some appropriate threshold in order to determine which hypothesis to accept. Since Λ is used by the decision rule only in comparison to a threshold—call it Λ_0—any one-to-one and monotonic transformation on Λ may be performed on both Λ and Λ_0, thereby generating a modified likelihood ratio to be compared with a modified threshold. In light of this, it is convenient to consider the log-likelihood ratio

$$\begin{aligned} \mathcal{L} &= \ln \Lambda \\ &= \sum_{j=1}^{J} \left[F_n(j) \ln\left(1 + \frac{S(j)}{\lambda^{(0)}}\right) - S(j) \right]. \end{aligned} \tag{9}$$

In its present form, the log-likelihood ratio detection statistic embodied in Eq. (9) has been derived under the implicit assumption that $\lambda^{(0)}$ is known. Since it is unlikely that $\lambda^{(0)}$ will be known *a priori*, we must now confront the problem of estimating $\lambda^{(0)}$. Intuitively, it should be expected that if the target aerosol signal is absent, then $\lambda^{(0)}$ can be estimated by averaging the count data $F_n(j)$ over the sample period index j. Strictly speaking, in the event that $\lambda^{(0)}$ is unknown, the likelihood ratio test requires that the parameter $\lambda^{(0)}$ be obtained as the maximum likelihood estimate under the null hypothesis H_0. In this regard, the necessary estimate is obtained by choosing $\lambda^{(0)}$ to maximize the probability of receiving the count data samples, viz., to maximize

$$p(\{F_n(j)\}|H_0) = \prod_{j=1}^{J} \frac{[\lambda^{(0)}(j)]^{F_n(j)}}{F_n(j)!} \exp(-\lambda^{(0)}(j)).$$

The maximum likelihood estimate is obtained by solving the equation $dp(\{F_n(j)\}|H_0)/d\lambda^{(0)} = 0$ and this process leads to

$$\frac{dp(\{F_n(j)\}|H_0)}{d\lambda^{(0)}} = -p(\{F_n(j)\}|H_0) \sum_{j=1}^{J} \left(\frac{F_n(j)}{\lambda^{(0)}} - 1 \right) = 0,$$

from which it follows

$$\lambda^{(0)} = \frac{1}{J} \sum_{j=1}^{J} F_n(j).$$

Note that this is simply the average of the count data $F_n(j)$ for the case when the signal aerosol is absent.

Choosing the detection threshold is just as important as computing the log-likelihood ratio. It is important to note that the choice of the threshold is intimately related to the evaluation of the performance of the detector, and to this purpose, the detection threshold $\mathcal{T}$ must be chosen to provide a given false alarm rate for the detector. This particular choice of the detection threshold provides the maximum probability of detection for a given probability of false alarm. It is important to note that the choice of the "correct" probability of false alarm is, of course, a question for the user of the detector to decide and one on which physical science has no bearing. However, it should be stressed that it is not possible to operate any detection system based on zero false alarm probability; it can be never be absolutely guaranteed that a zero probability of false alarm will occur for any physically realizable detector. The probability of false alarm, P_{FA}, is defined as

$$P_{\mathrm{FA}} = \Pr\{\mathcal{L} \geq \mathcal{T} | H_0\}$$

and the associated probability of detection, P_{D}, is given by

$$P_{\mathrm{D}} = \Pr\{\mathcal{L} \geq \mathcal{T} | H_1\}.$$

The calculation of the detection threshold $\mathcal{T}$ using the exact multivariate Poisson probability distribution is a difficult problem. However, for sufficiently long sampling times t_s, the count data $\{F_n(j)\}$, and even, more pronouncedly, the detection statistic $\mathcal{L}$ would tend to Gaussian variates. Hence, a Gaussian approximation for $\mathcal{L}$ allows $\mathcal{T}$ to be more easily computed. Applying the Lindelberg condition of the central limit theorem [4] to the sum of random variables in Eq. (9), it is straightforward to show that the log-likelihood ratio can be approximated by a Gaussian variate whose mean and variance are determined as follows. In the absence of a target aerosol signal (hypothesis H_0), the log-likelihood ratio $\mathcal{L}$ has a mean and variance given by

$$\mathbf{E}(\mathcal{L}|H_0) = \sum_{j=1}^{J} \left[\lambda^{(0)} \ln\left(1 + \frac{S(j)}{\lambda^{(0)}}\right) - S(j) \right]$$

and

$$\mathrm{Var}(\mathcal{L}|H_0) = \sum_{j=1}^{J} \ln^2\left(1 + \frac{S(j)}{\lambda^{(0)}}\right) \lambda^{(0)}.$$

Similarily, in the presence of the target aerosol signal (hypothesis H_1), the mean and variance of the log-likelihood ratio assumes the form

$$\mathbf{E}(\mathcal{L}|H_1) = \mathbf{E}(\mathcal{L}|H_0) + \sum_{j=1}^{J} S(j) \ln\left(1 + \frac{S(j)}{\lambda^{(0)}}\right)$$

and

$$\mathrm{Var}(\mathcal{L}|H_1) = \mathrm{Var}(\mathcal{L}|H_0) + \sum_{j=1}^{J} S(j) \ln^2\left(1 + \frac{S(j)}{\lambda^{(0)}}\right).$$

The approximate distribution of $\mathcal{L}$ under H_0 enables the determination of the threshold of any desired probability of false alarm. The approximate distribution of $\mathcal{L}$ under H_1 enables the determination of the probability of detection as a function of the signal-to-noise ratio. Consequently, in light of these relationships and the Gaussian approximation for the log-likelihood ratio $\mathcal{L}$, it is evident that to achieve a probability of false alarm, P_{FA}, of α requires the solution of

$$P_{\mathrm{FA}} = \alpha = \frac{1}{2} - \frac{1}{2}\mathrm{erf}(x/\sqrt{2}), \tag{10}$$

with the detection threshold calculated from

$$\mathcal{T} = \mathbf{E}[\mathcal{L}|H_0] + x\sqrt{\mathrm{Var}[\mathcal{L}|H_0]}. \tag{11}$$

Here, erf(·) denotes the error function. For this choice of detection threshold, the probability of detection can then be computed from

$$P_{\mathrm{D}} = \frac{1}{2} - \frac{1}{2}\mathrm{erf}(y/\sqrt{2}), \tag{12}$$

where

$$y \equiv \left(\mathcal{T} - \mathbf{E}[\mathcal{L}|H_1]\right) \Big/ \sqrt{\mathrm{Var}[\mathcal{L}|H_1]}. \tag{13}$$

For the likelihood ratio test that has been formulated, the performance of the detection statistic depends only on the dimensional parameter J, the Poisson parameter $\lambda^{(0)}$ associated with the background aerosol samples, and the signal-to-noise ratio (SNR) which is defined as

$$\mathrm{SNR} = \frac{1}{J} \sum_{j=1}^{J} \frac{S(j)}{\lambda^{(0)}}.$$

In examining the probability of detection, there are several different parameters that can be varied. In the following simulations, a Poisson parameter $\lambda^{(0)} = 10$ is used for the representation of the aerosol background noise level. In Figures 7–8, the probability of detection (P_{D}) is shown against the probability of false alarm (P_{FA}), parametrized by the SNR for $J = 10$. Recall that J is the number of aerosol samples used in the computation of the log-likelihood ratio. Figures 9–10 are analogous, but they correspond to the case for $J = 25$ aerosol samples used in the detection process. As would be expected, observe that the greater J is, the better the performance of the detection algorithm. These curves show the fact that for a fixed SNR, the detector with the most aerosol samples has the best detectability. Figures 11 and 12 depict the probability of detection versus the signal-to-noise ratio for several prescribed (i.e., fixed) probabilities of false alarm for $J = 10$ and 25, respectively. It is interesting to note that for SNRs greater than or equal to about 0.4, the detection performance curves suggest that P_{FA} depends only on the number of aerosol samples, J, used in the detection process and is independent of the SNR (viz., of the signal level of the target aerosol). Consequently, in this regime of operation, the detection threshold can be calculated without knowledge of the signal-to-noise ratio and need not be adjusted to keep a constant false alarm rate throughout the detection procedure. In this regime, the detection algorithm exhibits constant false alarm rate (CFAR) behavior.

In summary then, the log-likelihood ratio detection algorithm can be implemented as follows: (1) choose α for the desired P_{FA} using Eqs. (10) and (11); (2) compute the background Poisson parameter $\lambda^{(0)}$ using only samples obtained from the pure aerosol background; (3) compute the log-likelihood ratio $\mathcal{L}$ (cf. Eq. (9)); (4) calculate the detection threshold $\mathcal{T}$ from Eqs. (12) and (13); and, (5) compare $\mathcal{L}$ with $\mathcal{T}$ and declare the target (signal aerosol) to be present if $\mathcal{L} \geq \mathcal{T}$.

IV. SOME NUMERICAL EXAMPLES

In this section, the use of the detection procedure based on the log-likelihood ratio statistic, is illustrated by some numerical examples. The background aerodynamic particle size distribution function is assumed to be represented by a gamma size distribution with the shape parameter $\alpha = 1.5$ and the scale parameter $\beta = 2.5\ \mu$m. In the following examples, the total number of aerosol particles in the aerodynamic diameter range $[D_1, D_2] = [1.0, 15.0]\ \mu$m in the sampled volume V_s, is used as the count data for the construction of the log-likelihood ratio detection statistic.

Example 1: Figure 13 shows a target aerosol signal in the form of the time history of the number concentration at a receptor position where an APS has been deployed. In this example, the peak amplitude of the number concentration is 40 cm^{-3}—the time of passage of the quasi-instantaneous puff over the receptor position is about 15 min. For this example, the particle number density N^* of the gamma size distribution, used to model the background particle size distribution, was adjusted to yield a background aerosol count with an intensity $\lambda^{(0)} = 10.0$. A sampling time of $t_s = 60$ s was chosen for the APS detection system operating at the aerosol flowrate $Q_s = 83.3$ $cm^3\ s^{-1}$—this sampling time (i.e., exposure time) was chosen to ensure a statistical sampling error of smaller than or equal to 0.5 percent. The detection performance will be expressed in terms of the probability of detection, P_D, as a function of the number of aerosol samples, J, used in the detection scheme, for a fixed probability of false alarm P_{FA}. For this example, $P_{FA} = 10^{-4}, 10^{-3}$, and 10^{-2} were chosen. It is assumed that aerosol samples used in the detection procedure were collected immediately after the initial arrival of the transient target aerosol signal. The performance of the detector is shown in Figure 14. Observe that the probability of detection $P_D \approx 1.0$ for $J \geq 7$ and that the detection probability is independent of P_{FA} after this point. With the prescribed sampling time of $t_s = 60$ s, this implies that the APS was able to detect the target signal with certainty approximately 7 minutes after the initial arrival of the transient.

Example 2: The experiment described in Example 1 was repeated, but this time the target aerosol signal possesses the form depicted in Figure 15. The target signal is a broader, semi-continuous puff with a peak amplitude of approximately 1.4 cm^{-3}—the time of passage of this transient signal over the receptor position is about 150 min. For this example, the particle number density N^* was adjusted to provide a background intensity of $\lambda^{(0)} = 5.0$. It is noted that the background aerosol clutter level is lower than that assumed in Example 1. A sampling time $t_s = 20$ s was chosen for the APS, operating at the volume flowrate of $Q_s = 83.3$ $cm^3\ s^{-1}$ in order to yield a prescribed value for the relative fluctuation intensity of the total number of aerosol particles

in the sampled volume of 1 percent (viz., this means that one can expect a standard deviation in the total number of particles of approximately 1 percent of the expected total number due to Poisson distributed fluctuations of the number of particles sampled about their mean gamma size distribution). For this example, the number of samples used for the detection procedure is $J = 45$ and, with a sampling time of $t_s = 20$ s, this implies that a detection decision is made after every sampling duration of 15 min. Figure 16 shows that detection probability, at prescribed false alarm probabilities of 10^{-3}, 10^{-2} and 10^{-1}, for each sampling duration of 15 min after the initial arrival of the target signal. In the plot, the time index along the abscissa corresponds to the number index of each detection event of sampling duration 15 min. In this case, the detection performance is fair since for $P_{FA} = 0.01$, there is a 0.87 probability of detection of the target signal (shown in Figure 15) for the detection period that includes the peak of the target signal.

V. CONCLUSIONS

In this paper, the basic theory and algorithms for a statistical sampling error model and a detection scheme for the aerodynamic particle size analyzer, has been developed. It has been shown how the Poisson process can be used as the basis for a sampling error model which characterizes the fluctuations of a particle size distribution-related variable due to stochastic fluctuations of the number of particles sampled in each particle size interval. The model developed here allows for the proper specification of the required sampling program for the APS in order that the sampling error be below some fixed error level. With the prescription of an appropriate sampling program, this paper presents an algorithm for detecting a target aerosol signal against a clutter aerosol background. The detection algorithm is based on the generalized likelihood ratio test and can be applied to the detection of an aerosol signal sequence in a set of J aerosol samples with a common aerosol clutter background. It has been shown how the probability of detection of the detector can be determined for a prescribed probability of false alarm, including the effects of the signal-to-noise ratio, the intensity of the background aerosol clutter, and the number of aerosol samples, J, used in the detection process. Some detection results have been presented which illustrate the wide variations in performance which may be observed, dependent on the background clutter condition, the form of the target aerosol signal to be detected, the choice of the detection threshold, and the signal-to-noise ratio.

Practical application of the detection algorithm described in this paper must also consider the issues related to the operational characteristics of the APS. One issue concerns the efficiency

of the APS, viz., what fraction of the particles in the sampled volume are actually counted by the instrument. Another issue is the robustness of the APS and the concomitant failure rate of the instrument when employed in a continuous field sampling mode for monitoring and detection. Finally, the paper has only addressed the detection performance of a single APS operating as a centralized system—the design and performance of a distributed signal detection system based on an array of aerodynamic particle size analyzers which incorporates the possibility of sensor failure needs to be addressed.

ACKNOWLEDGMENTS

The author would like to thank Mr. Stan Mellsen and Mr. George Wellon for their valuable comments on the paper and the group leader, Dr. S. J. Armour, for her help and encourgement.

REFERENCES

1. Blackford, D., Quant, F. and Sem, G., "An Improved Aerodynamic Particle Sizer", Fine Particle Society Annual Meeting, Boston, 1987.

2. Bharucha-Reid, A. T., *Elements of the Theory of Markov Processes and Their Applications*, New York: McGraw-Hill, 1960.

3. Bencala, K. E., Melice, J. L. and Demuth, C., "Statistical Distribution of Daily and High Atmospheric SO_2 Concentrations", Atmospheric Environment, Vol. 6, 941–950, 1982.

4. Feller, W., *An Introduction to Probability and its Applications*, New York: Wiley, 1971.

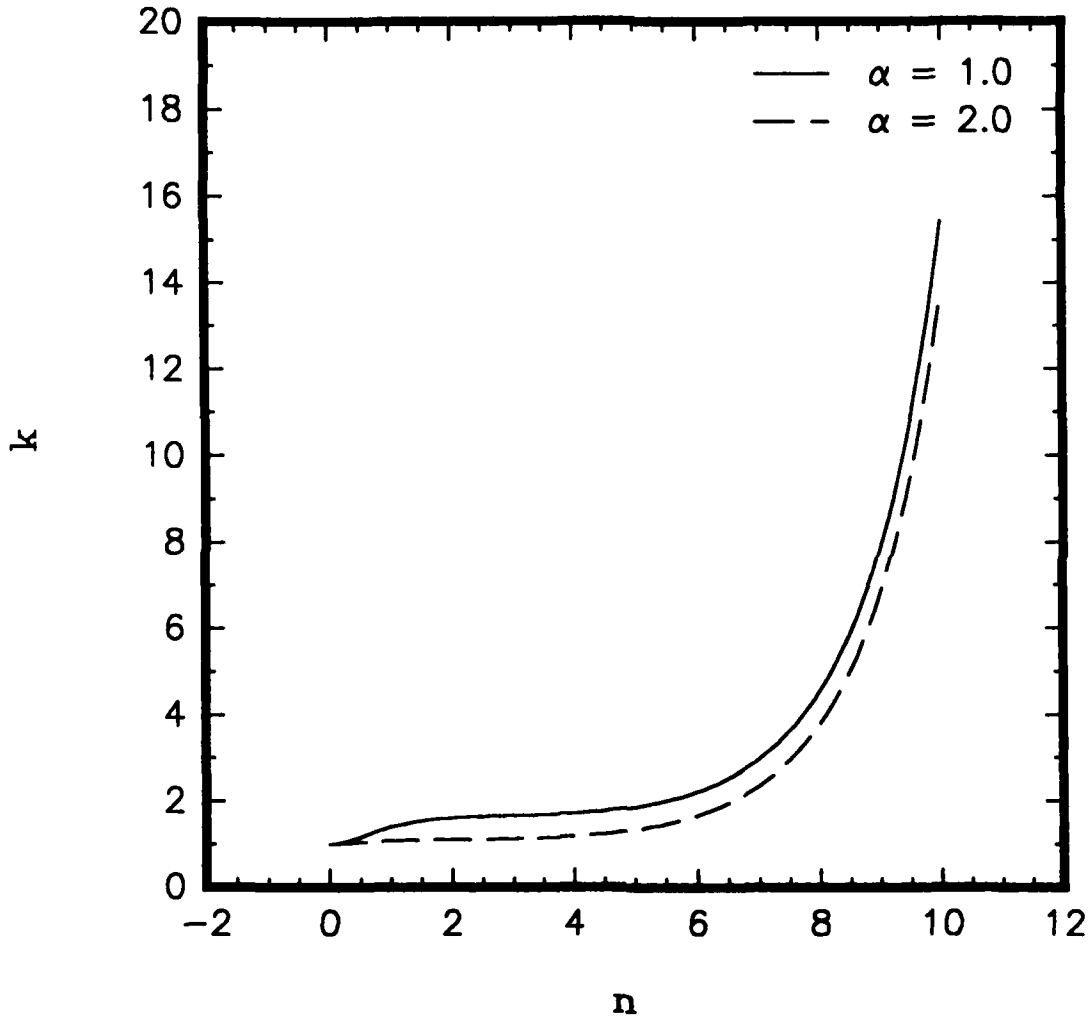

FIGURE 1

A plot of $k \equiv 1/\kappa = \langle N \rangle / N_e$ as a function of the aerodynamic diameter exponent n for $\alpha = 1.0$, 2.0 and $\beta = 1.0$ μm.

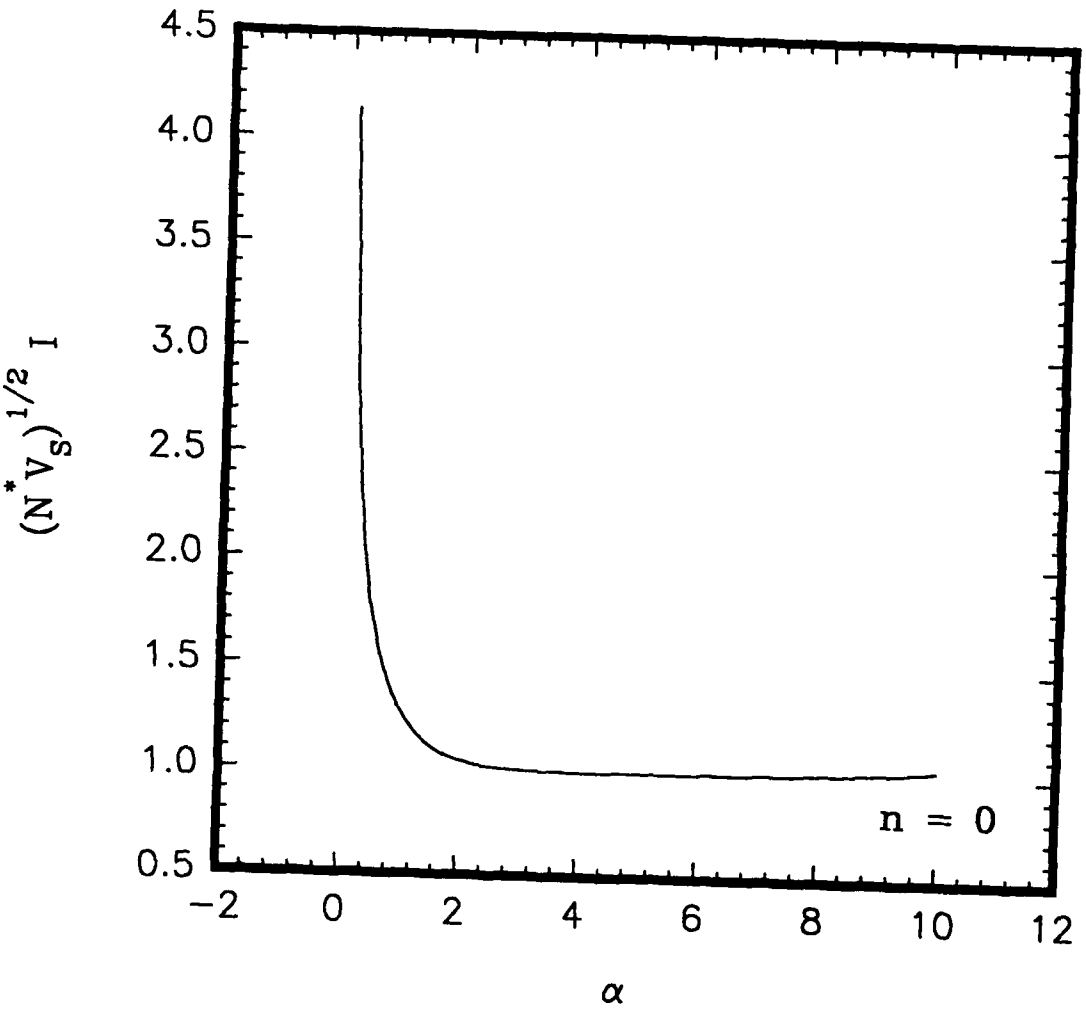

FIGURE 2

Curve of $(N^* V_s)^{1/2} I$ as a function of the gamma size distribution shape parameter α for $\beta = 1.0$ μm and $n = 0$.

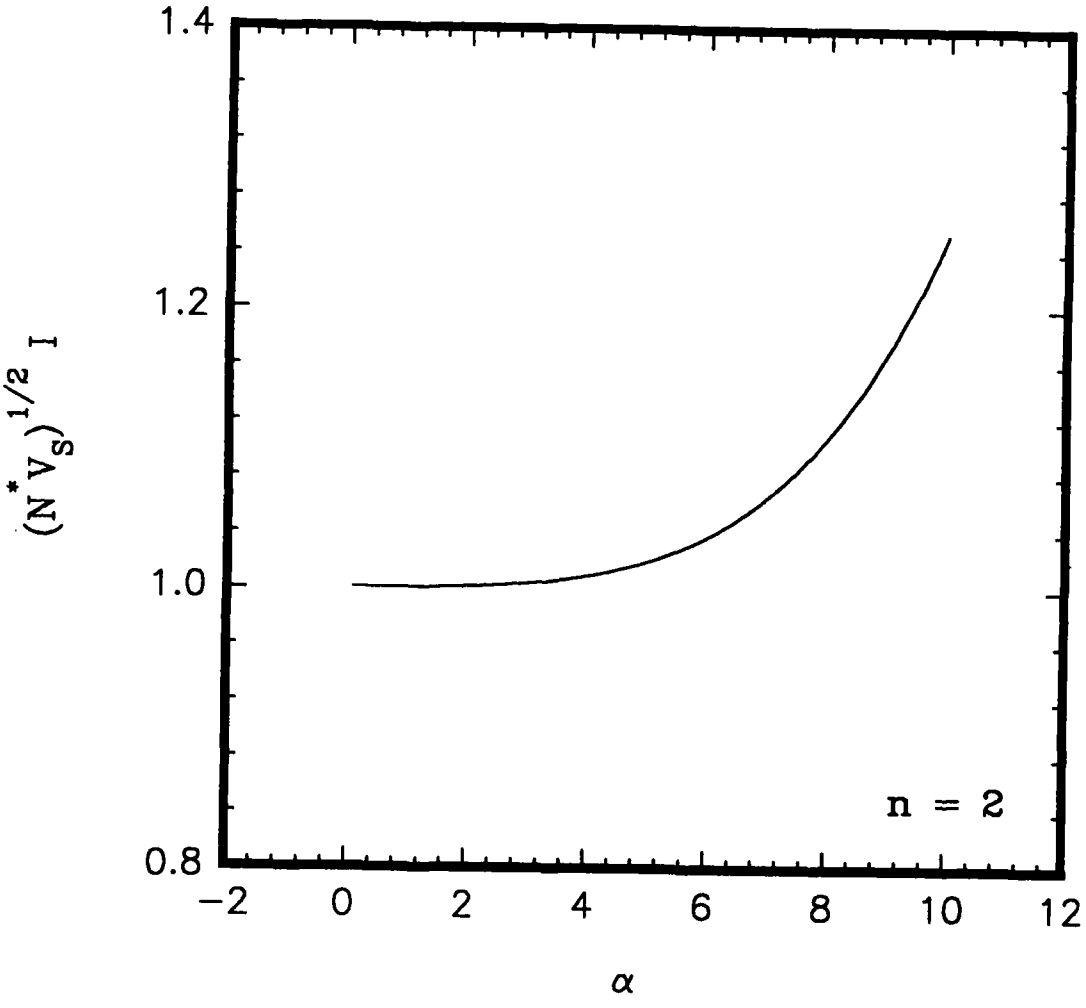

FIGURE 3

Curve of $(N^* V_s)^{1/2} I$ as a function of the gamma size distribution shape parameter α for $\beta = 1.0$ μm and $n = 2$.

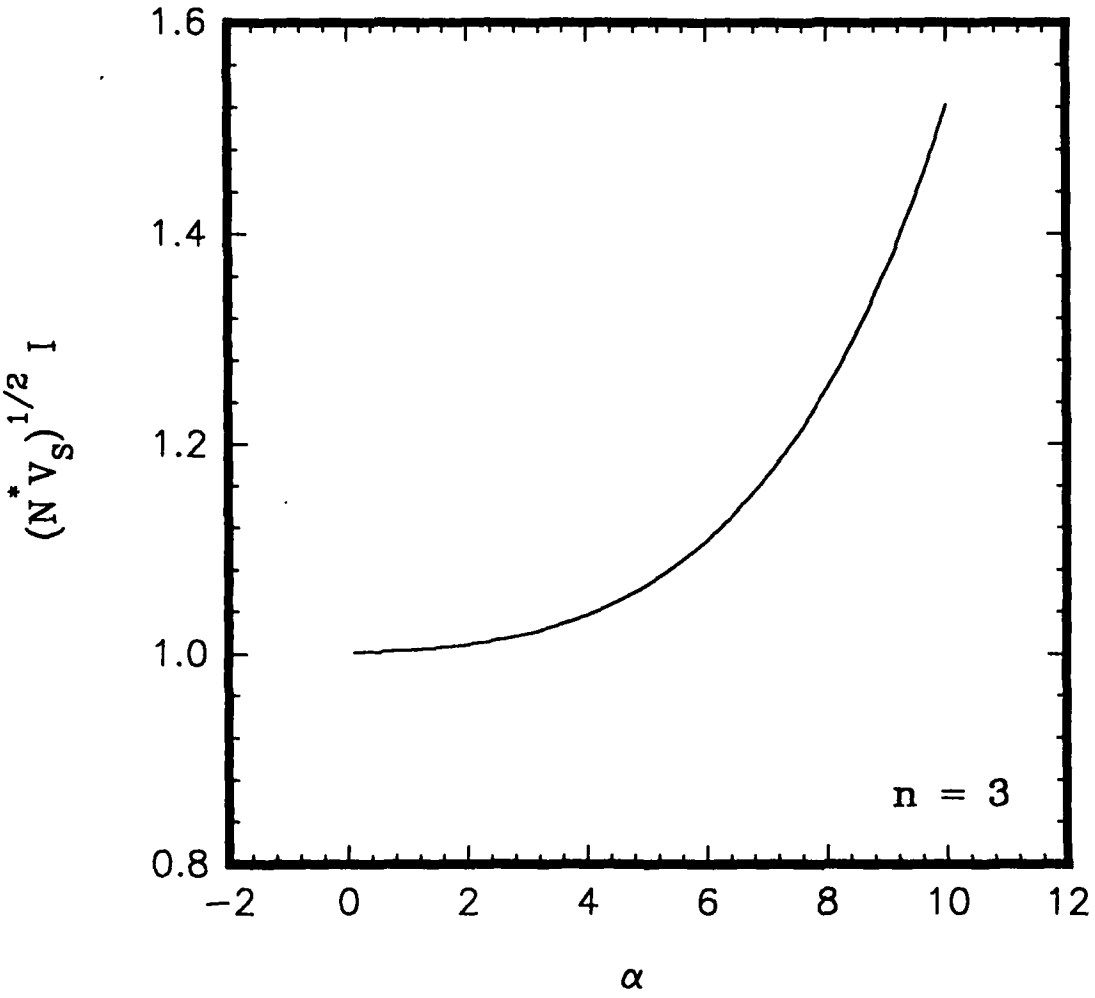

FIGURE 4

Curve of $(N^* V_s)^{1/2} I$ as a function of the gamma size distribution shape parameter α for $\beta = 1.0$ μm and $n = 3$.

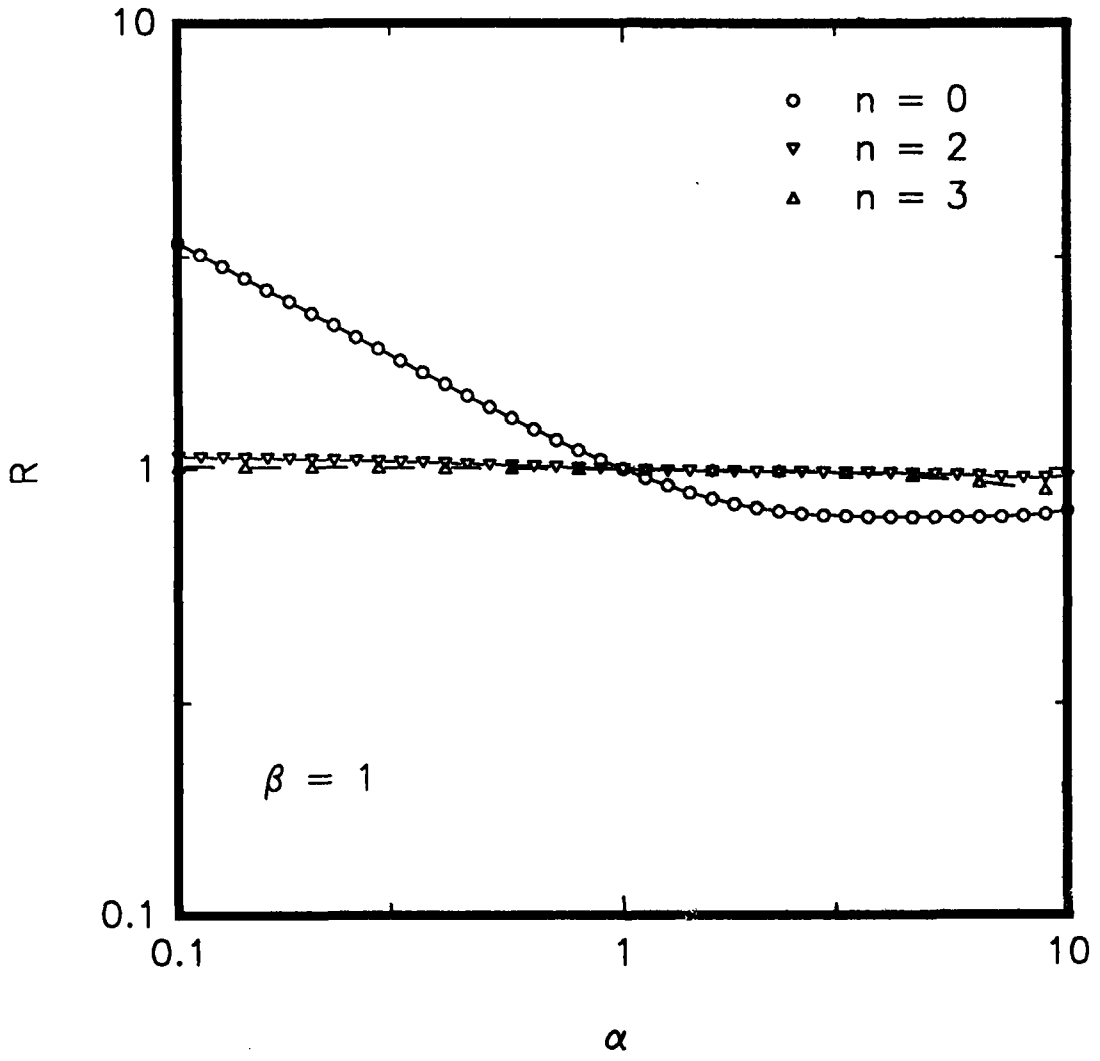

FIGURE 5

Ratio R as a function of the gamma shape parameter α for $n = 0, 2, 3$ and $\beta = 1.0\ \mu$m.

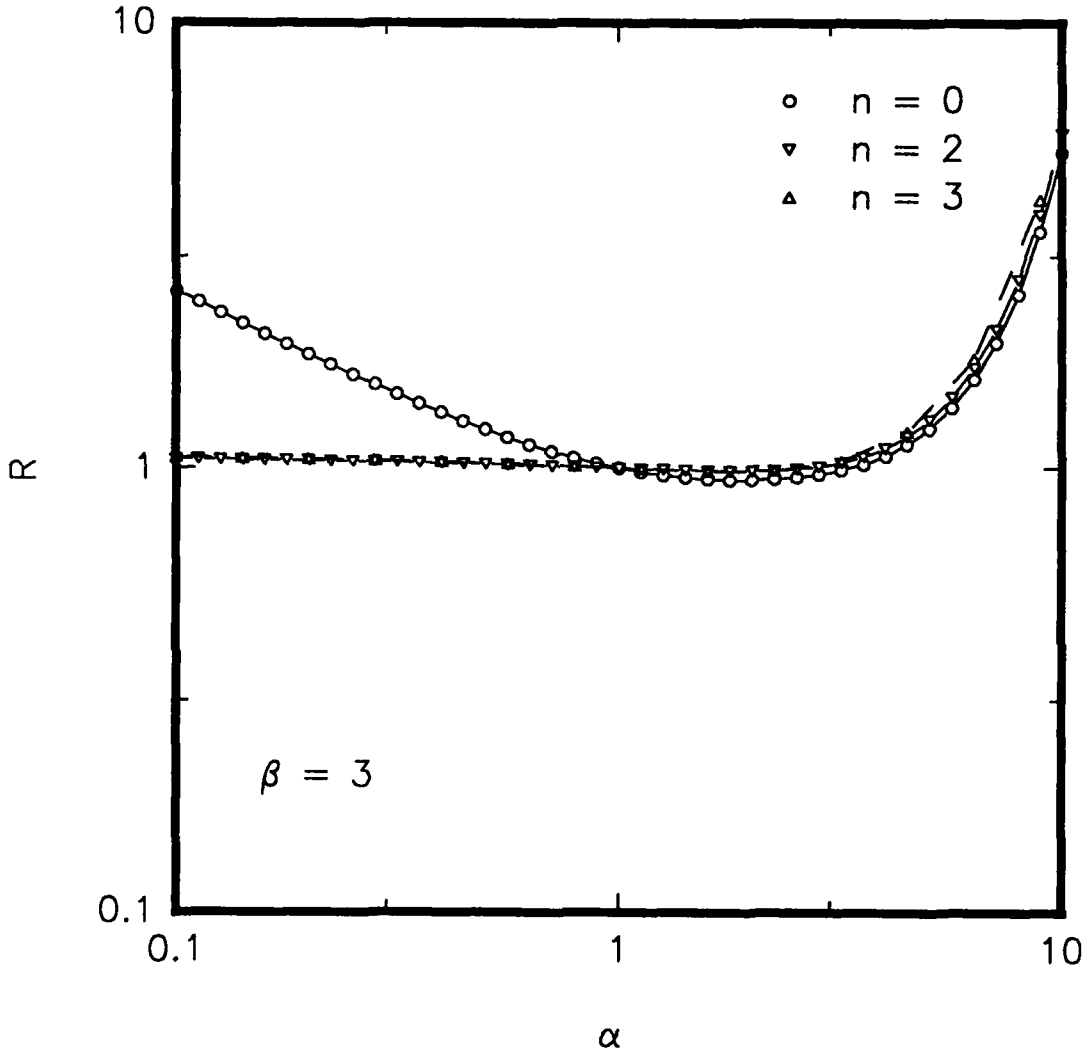

FIGURE 6

Ratio R as a function of the gamma shape parameter α for $n = 0$, 2, 3 and $\beta = 3.0$ μm.

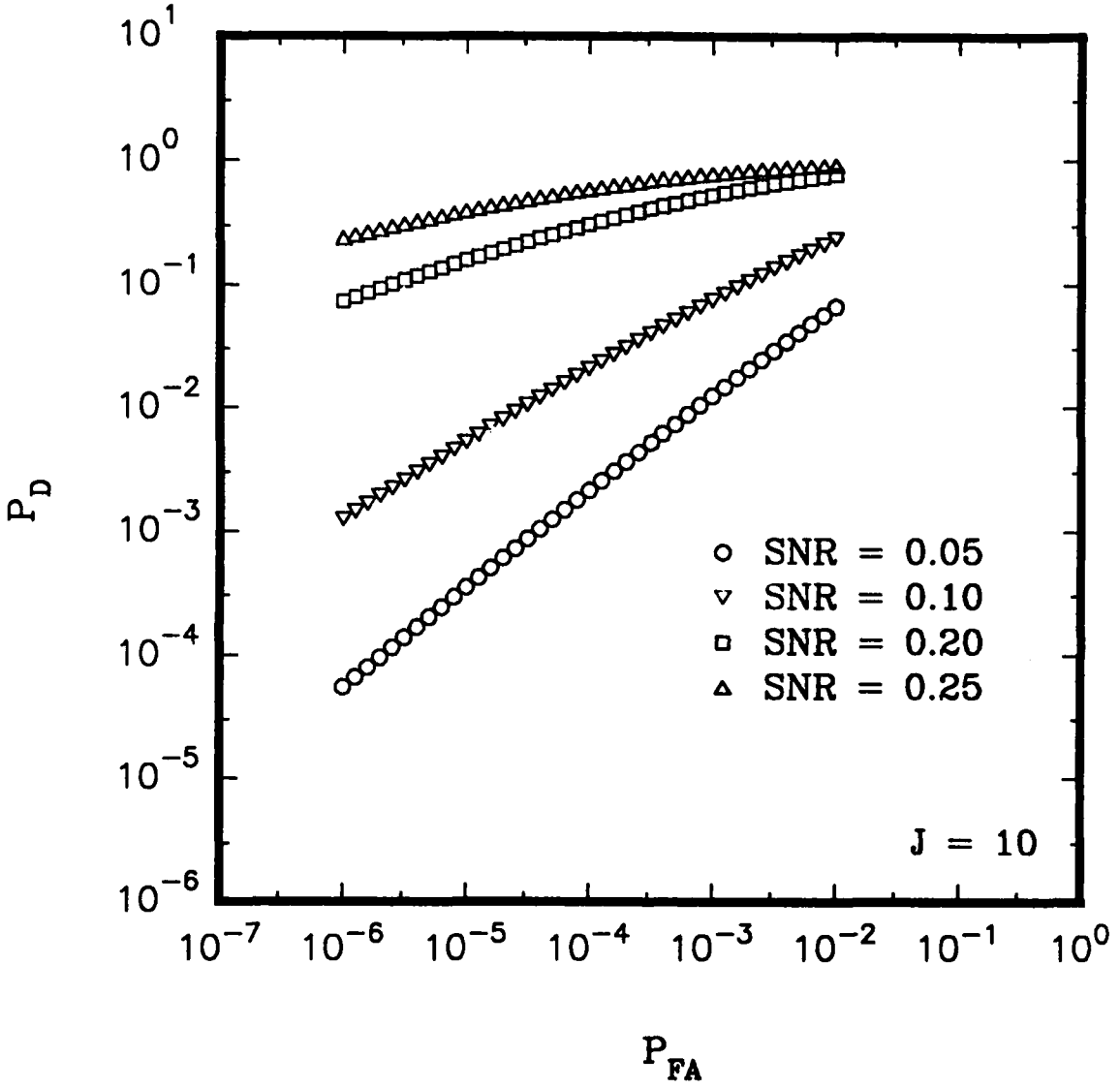

FIGURE 7

Probability of detection (P_D) versus probability of false alarm (P_{FA}), shown for several signal-to-noise ratio (SNR) values, given $J = 10$.

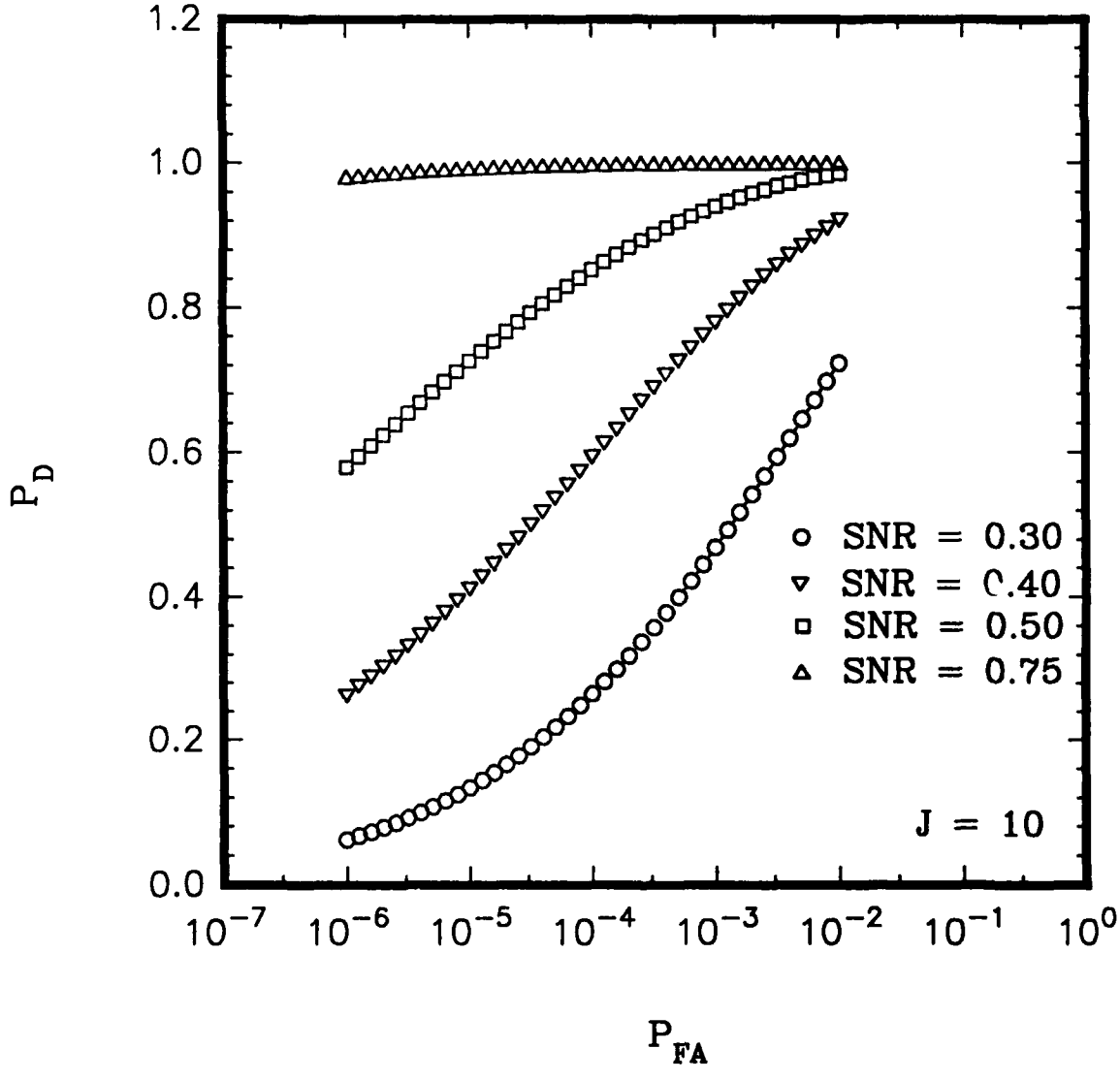

FIGURE 8

Probability of detection (P_D) versus probability of false alarm (P_{FA}), shown for several signal-to-noise ratio (SNR) values, given $J = 10$.

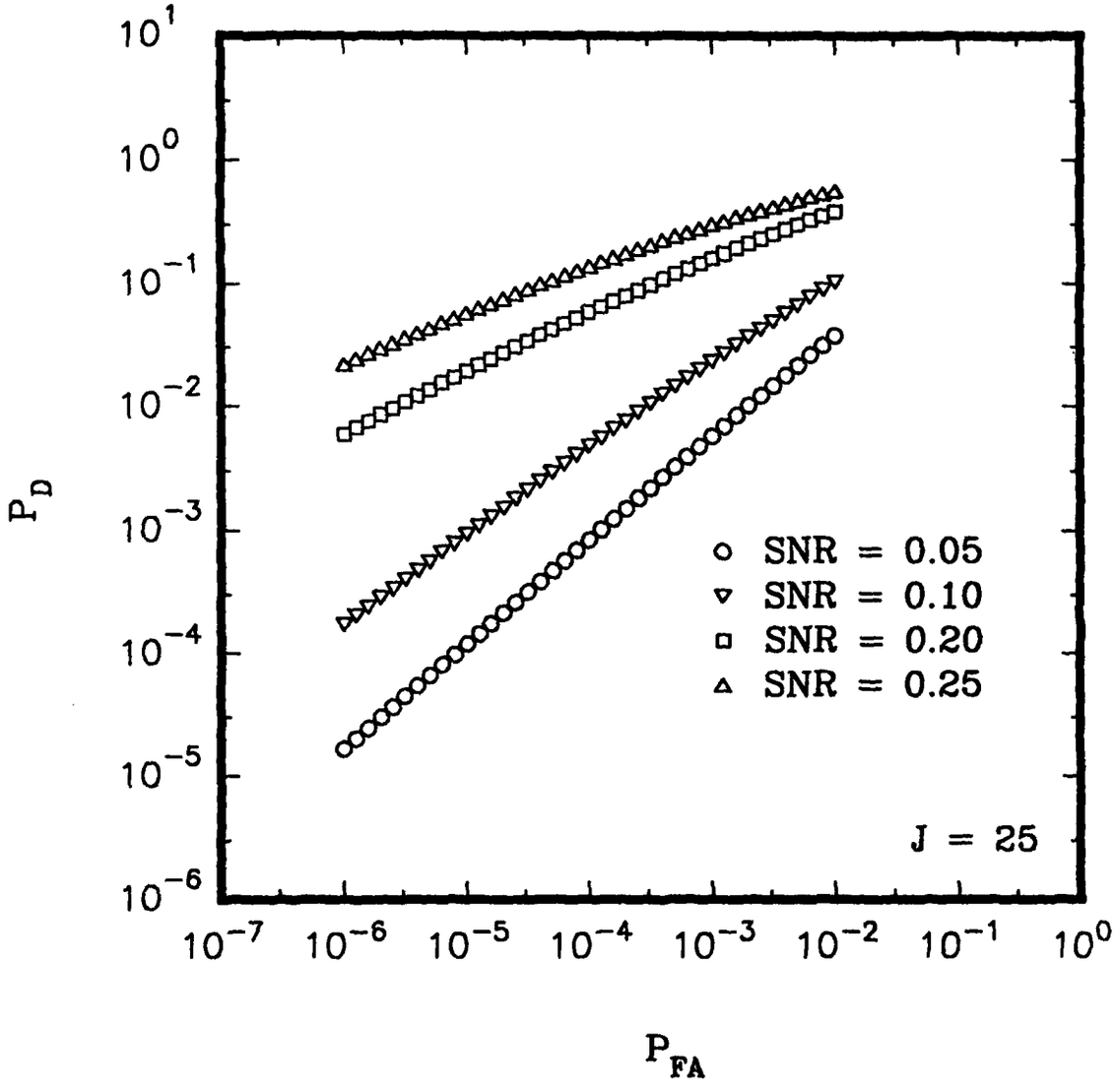

FIGURE 9

Probability of detection (P_D) versus probability of false alarm (P_{FA}), shown for several signal-to-noise ratio (SNR) values, given $J = 25$.

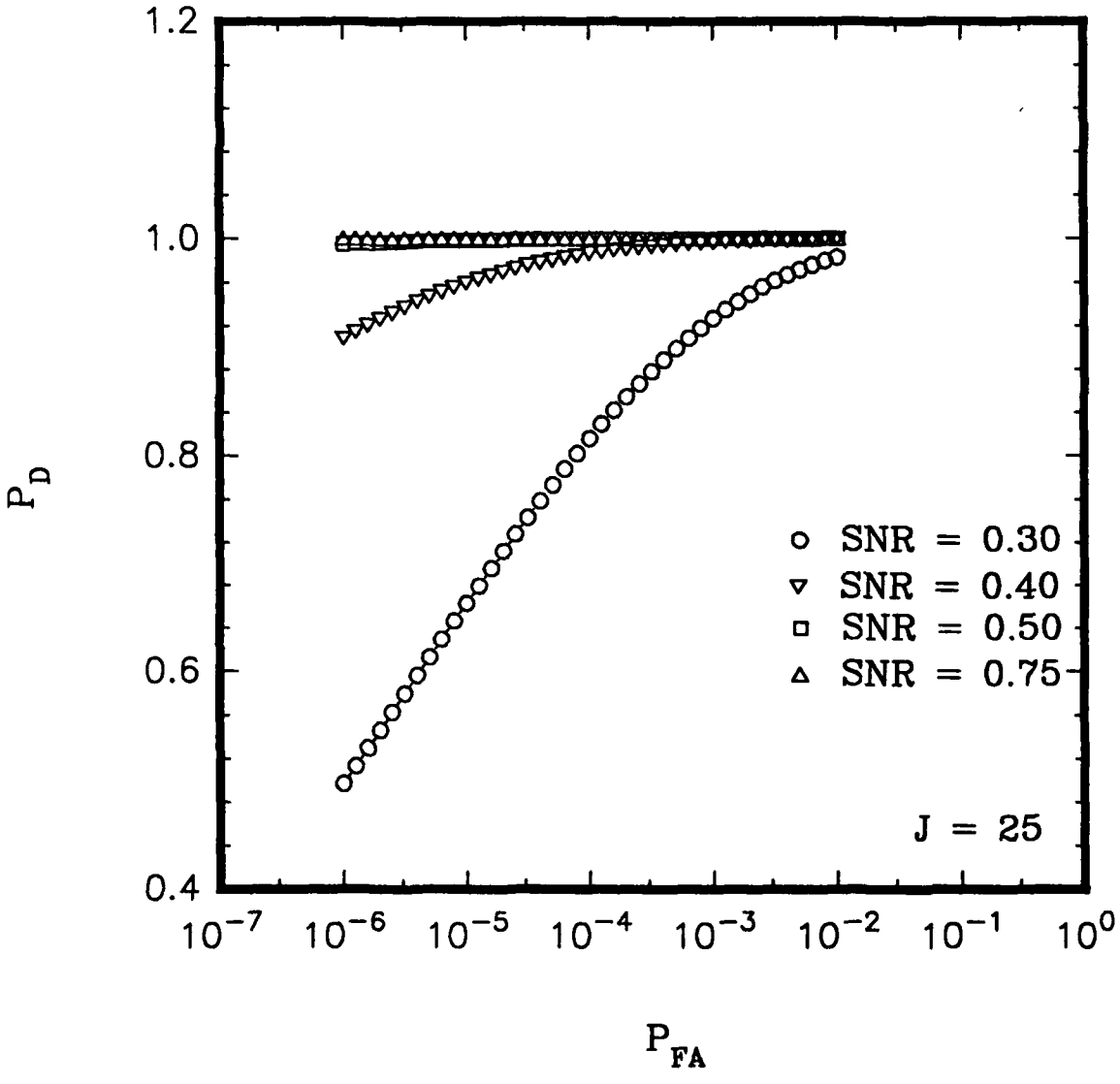

FIGURE 10

Probability of detection (P_D) versus probability of false alarm (P_{FA}), shown for several signal-to-noise ratio (SNR) values, given $J = 25$.

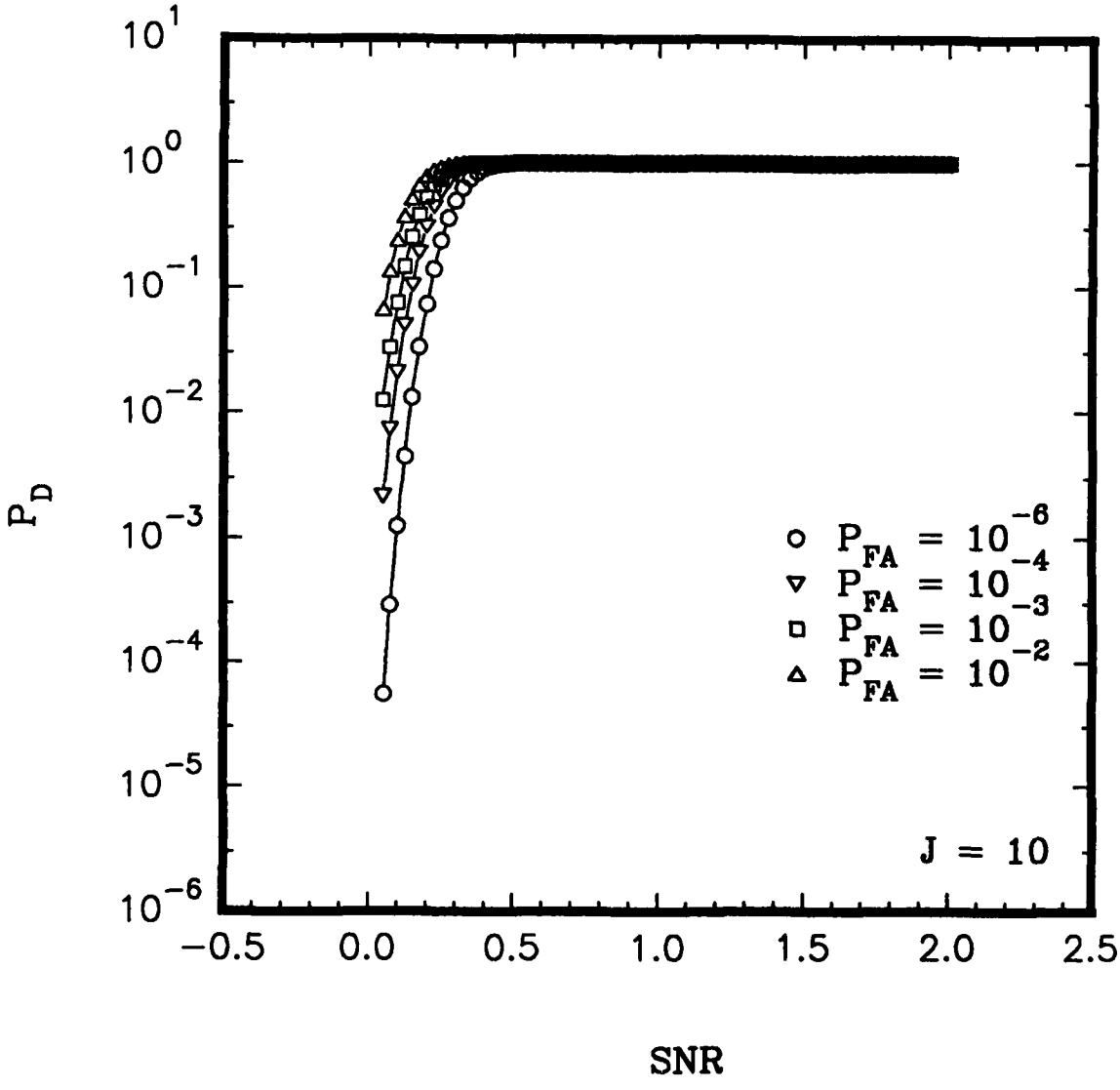

FIGURE 11

Probability of detection (P_D) versus signal-to-noise ratio (SNR) for $P_{FA} = 10^{-6}, 10^{-4}, 10^{-3}, 10^{-2}$ and $J = 10$.

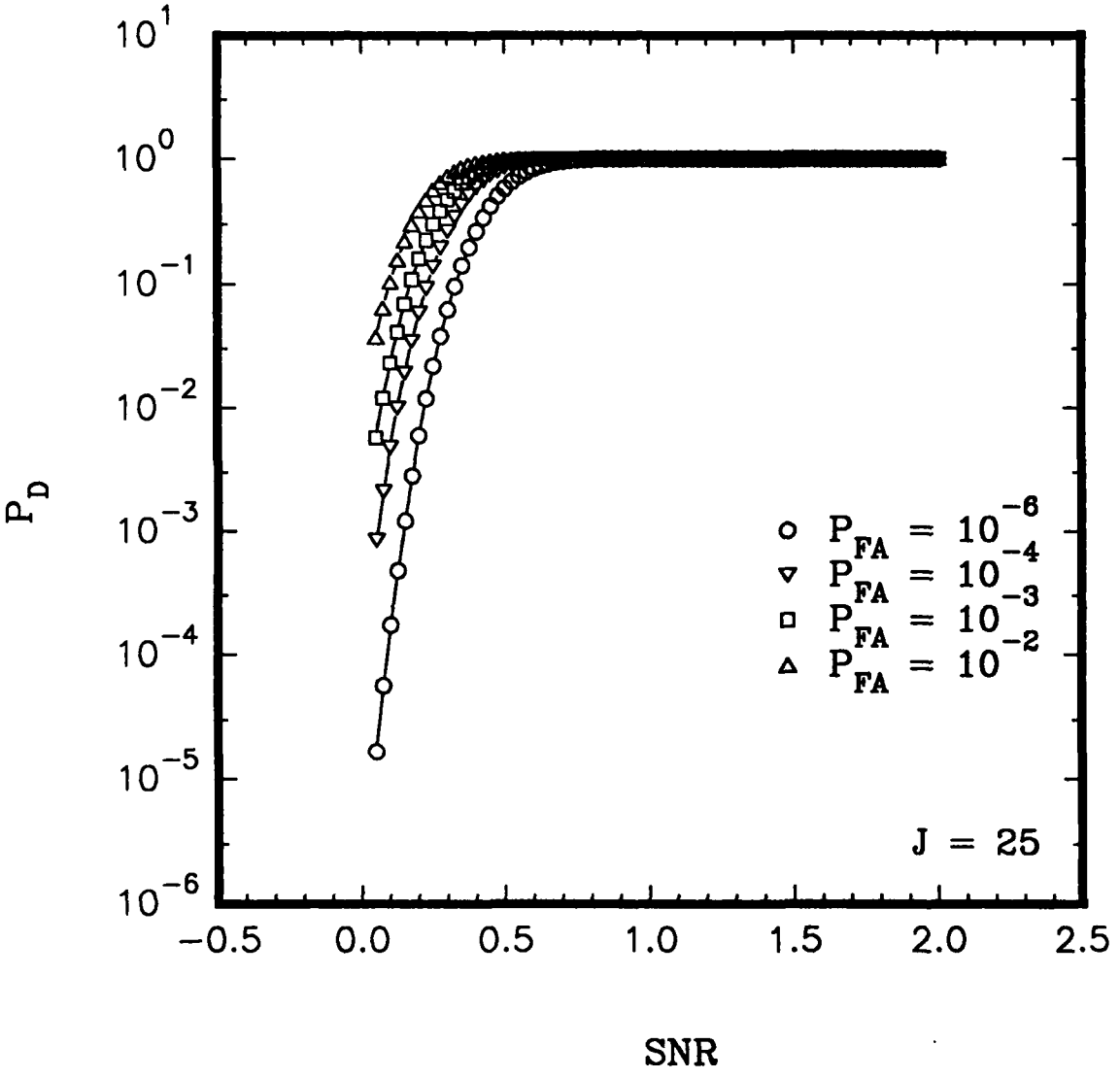

FIGURE 12

Probability of detection (P_D) versus signal-to-noise ratio (SNR) for $P_{FA} = 10^{-6}, 10^{-4}, 10^{-3}, 10^{-2}$ and $J = 25$.

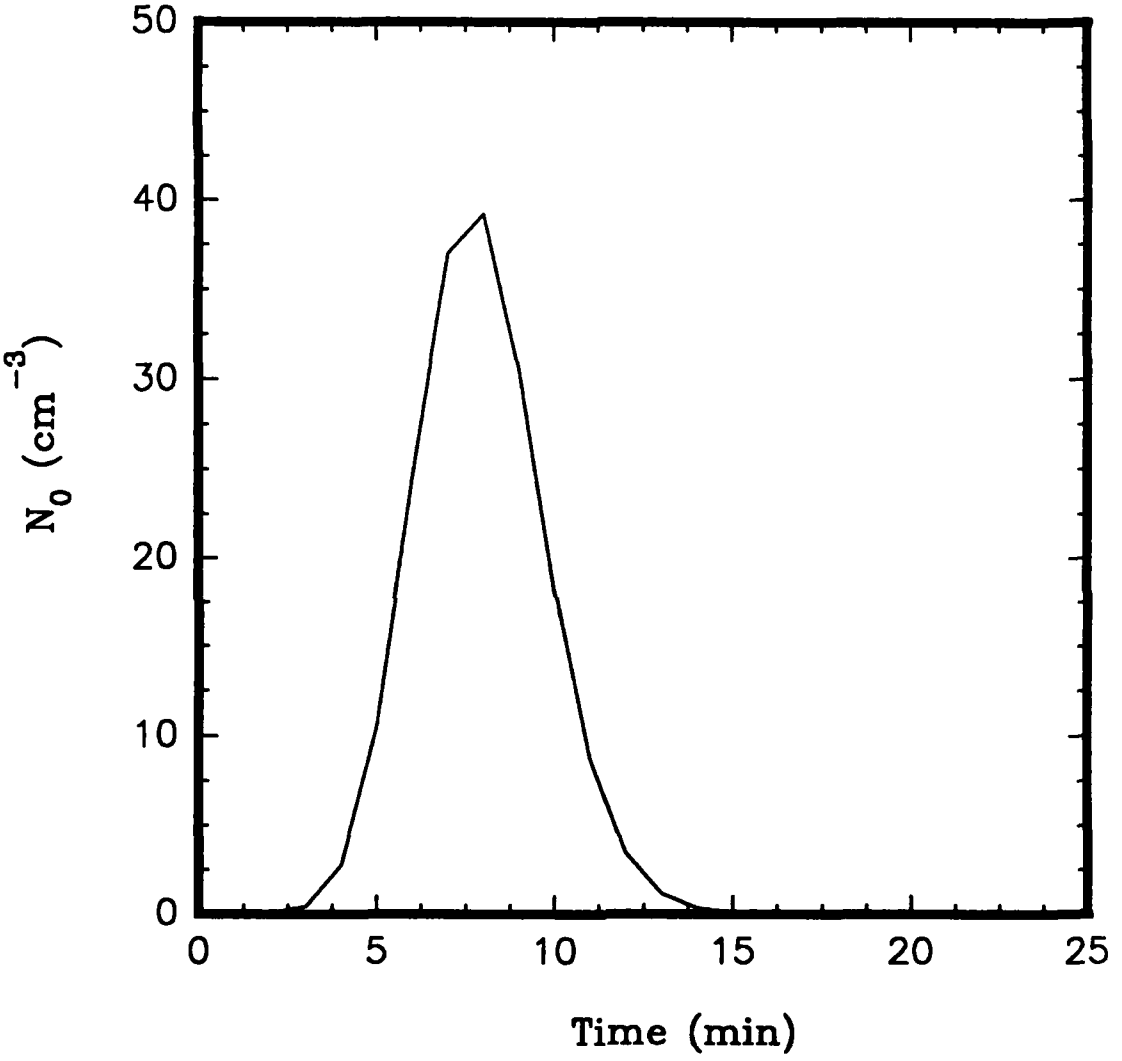

FIGURE 13

Time history of the number concentration of the target aerosol signal used in Example 1.

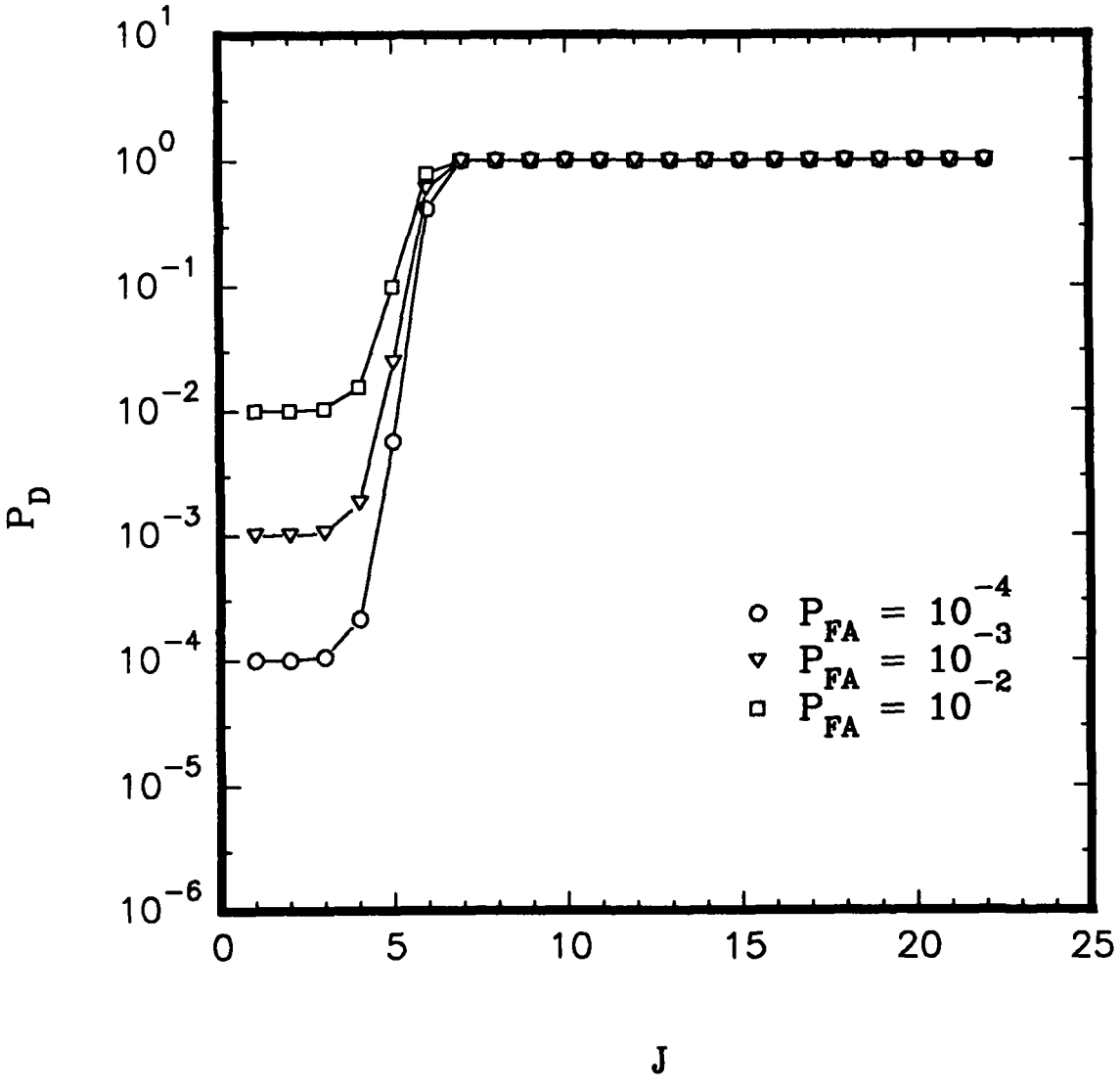

FIGURE 14

Probability of detection (P_D) for the transient aerosol signal depicted in Fig. 13 as a function of the number of aerosol samples, J, used in the detection process.

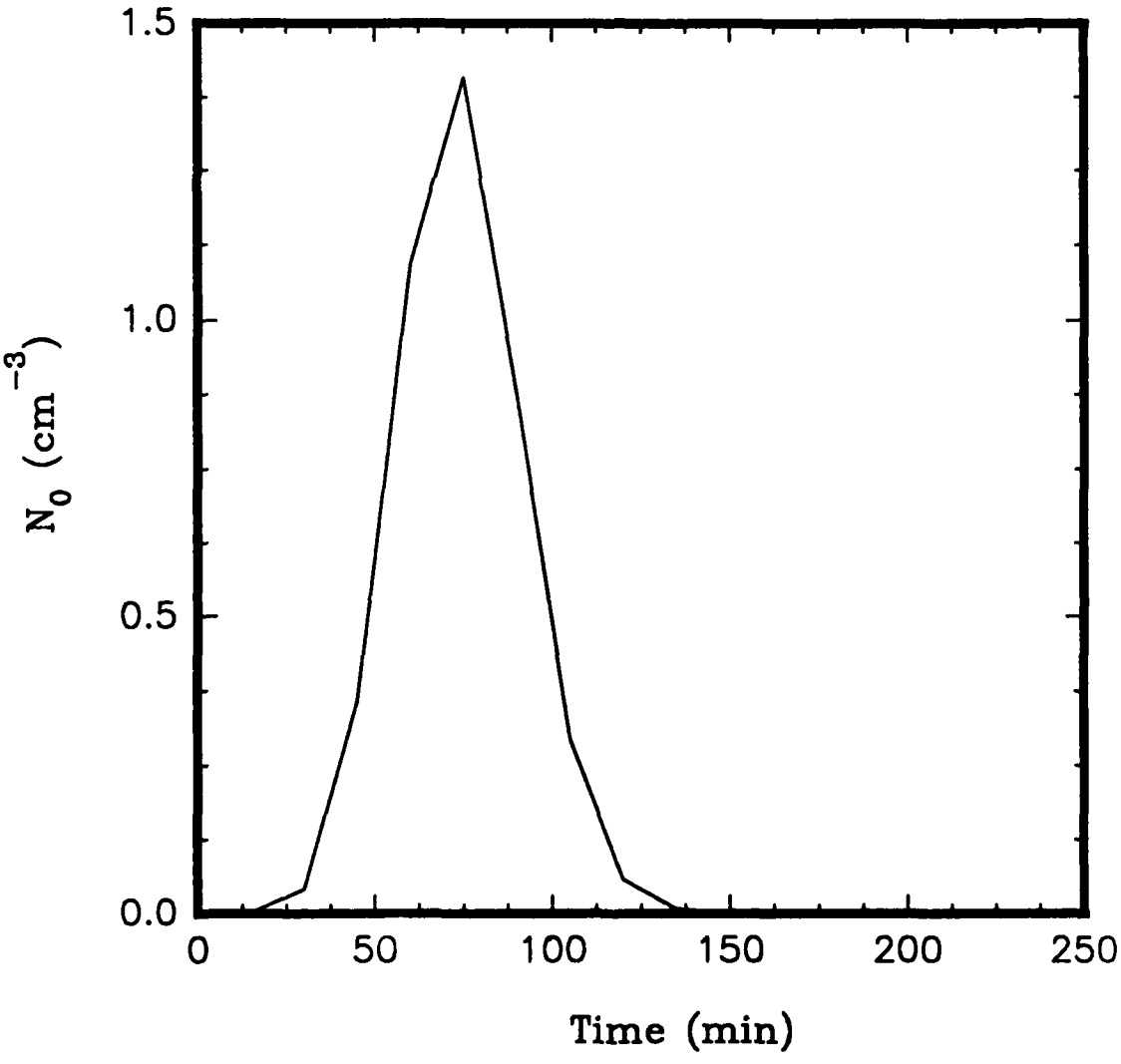

FIGURE 15

Time history of the number concentration of the target aerosol signal used in Example 2.

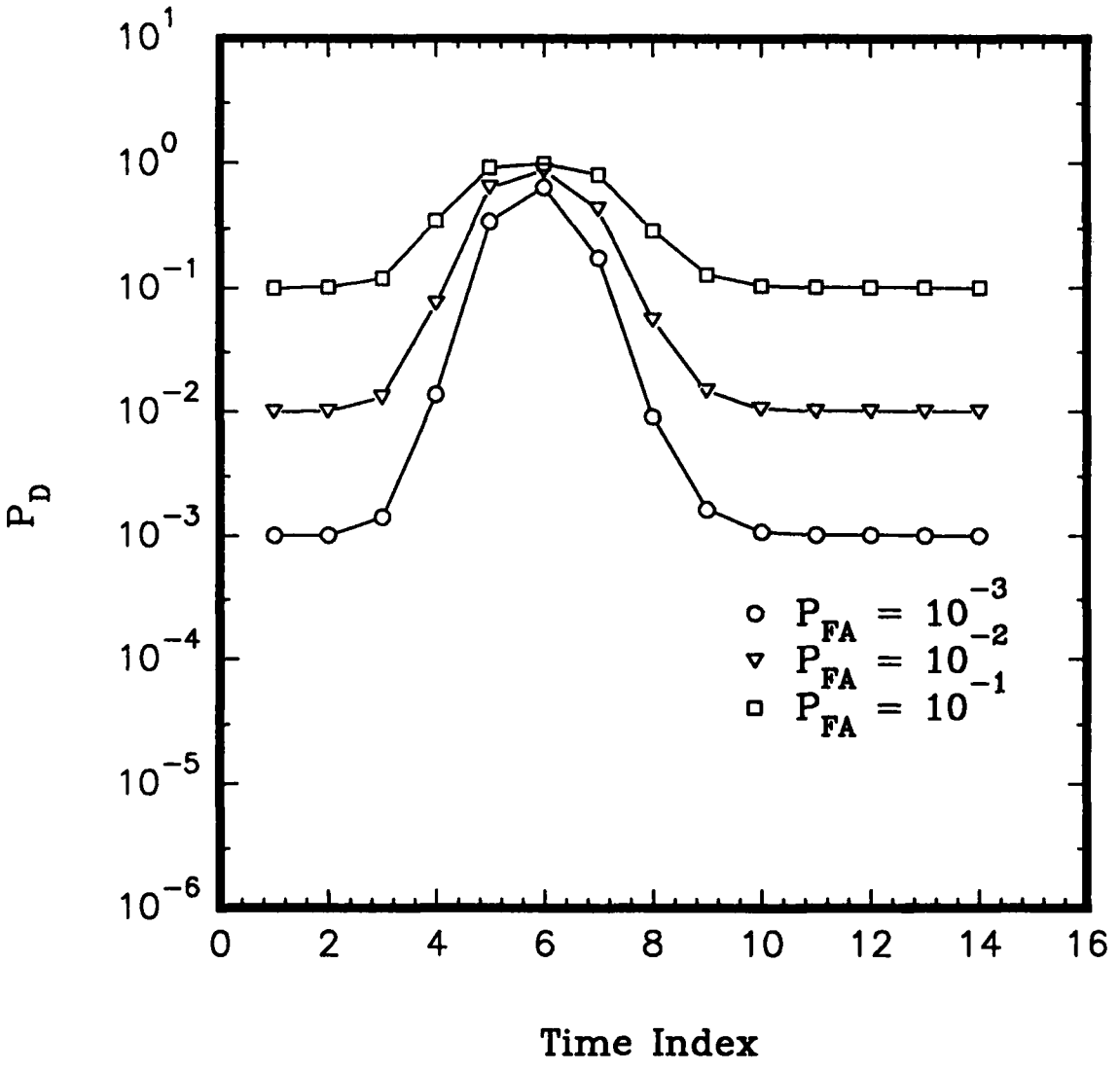

FIGURE 16

Probability of detection (P_D) for the transient aerosol signal depicted in Fig. 15 as a function of the detection period (indicated by the time index). Each detection period consisted of a sampling duration of 15 min.

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

Defence Research Establishment Suffield
Box 4000, Medicine Hat, AB. T1A 8K6

2. SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable)

Unclassified

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.)

A Formulation of a Stochastic Sampling Error Model and a Signal Detection Algorithm For The Aerodynamic Particle Size Analyzer (U)

4. AUTHORS (Last name, first name, middle initial. If military, show rank, e.g. Doe, Maj. John E.)

YEE, Eugene

5. DATE OF PUBLICATION (month and year of publication of document)

June 1991

6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)

33

6b. NO. OF REFS (total cited in document)

4

6. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

SR 549 (Final)

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

Defence Research Establishment Suffield, Box 4000, Medicine Hat, Alberta T1A 8K6

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)

PCN No. 351SP

9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)

SR 549

10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

Nil

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both offical languages unless the text is bilingual).

The determination of the distribution of airborne toxic particles as a function of the aerodynamic diameter provides important information as well as criteria for the definition of hazard as applied to levels of airborne contamination. This is because the aerodynamic particle size distribution embodies the information related to particle density, diameter, shape factor and slip correction that is critical for the characterization of particle motion in settling and impaction and it is these motions that are responsible for particle deposition in the respiratory tract and particle collection in aerosol sampling devices. For a given definition of hazard based on some parameter related to the aerodynamic size distribution, this paper develops a statistical sampling error model for the parameter that is based on the Poisson process. Given that an appropriate sampling program has been designed for the measurement of the size distribution-related parameter with the aerodynamic particle size analyzer, this paper proceeds to the derivation of an optimum detection algorithm for the detection of a signal aerosol sequence in a set of J aerosol samples with a common background. The detection algorithm is based on the generalized likelihood ratio test in which the received count associated with the aerosol sample is modeled as a Poisson distributed random variable. Performance analyses of the resulting algorithm, based on the probability of detection (P_D) versus signal-to-noise ratio for several given fixed false alarm probabilities (P_{FA}), are presented.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Aerodynamaic particle sizer
Poisson Process
Sampling error
Detection algorithm